JOURNAL OF BACTERIOLOGY, June 2011, p. 3150–3151 Vol. 193, No. 12
0021-9193/11/$12.00 doi:10.1128/JB.00344-11

Genome Sequence of the Spinosyns-Producing Bacterium *Saccharopolyspora spinosa* NRRL 18395▽

Yuanlong Pan,[1,2] Xi Yang,[1,2] Jing Li,[1] Ruifen Zhang,[1] Yongfei Hu,[1] Yuguang Zhou,[3] Jun Wang,[4] and Baoli Zhu[1]*

Microbial Genome Research Center, CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, P. R. China[1]; Graduate University of Chinese Academy of Sciences, CAS, Beijing 100049, P. R. China[2]; China General Microbiological Culture Collection Center, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, P. R. China[3]; and Beijing Genomics Institute, Shenzhen 518000, P. R. China[4]

Received 11 March 2011/Accepted 31 March 2011

***Saccharopolyspora spinosa* is a Gram-positive bacterium that produces spinosad, a well-known biodegradable insecticide that is used for agricultural pest control and has an excellent environmental and mammalian toxicological profile. Here, we present the first draft genome sequence of the type strain *Saccharopolyspora spinosa* NRRL 18395, which consists of 22 scaffolds.**

Saccharopolyspora spinosa belongs to the genus *Saccharopolyspora*, originally isolated from a soil sample collected from the Caribbean islands in 1982 (12). *S. spinosa* is the only bacterium that secretes spinosad, a natural pesticide consisting of spinosyn A and spinosyn D that kills most pests and has no toxic effect on humans (3, 14, 15). Spinosad is produced by Dow AgroSciences, which won the Presidential Green Chemistry Challenge Award for Designing Greener Chemicals in 1999, and is classified by the U.S. Environmental Protection Agency as an environmentally and toxicologically reduced risk material.

The whole-genome sequencing of *S. spinosa* will help scientists to better understand the regulation of the spinosyn biosynthesis pathway and its relationship to other genes, which will in turn allow more sophisticated strategies to improve the production of spinosad. According to the Genomes OnLine Database (GOLD) (10), genome sequencing of this organism has not been done yet, and there is only one completed and one draft sequence available for the genus *Saccharopolyspora*, specifically, that of *Saccharopolyspora erythraea*.

Here we present the whole-genome sequence of *Saccharopolyspora spinosa* NRRL 18395 (= CGMCC 4.1365), obtained using a whole-genome shotgun strategy with an Illumina genome analyzer (totaling ~2,906.97 Mb; ~330-fold coverage of the genome) and by Roche 454 pyrosequencing (totaling ~202.28 Mb; ~23.5-fold coverage of the genome) at the Beijing Genomics Institute.

The Illumina pair-end reads were assembled by SOAPdenovo (9), while 454 reads were assembled using Newbler Assembler software. The two assembled scaffolds were combined by sequence overlapping. We also filled hundreds of gaps by PCR followed by sequencing of the PCR product with the ABI 3730 platform, producing a final draft assembly of 22 scaffolds.

The annotation was done by using MetaGeneAnnotator (13) to predict the open reading frame, by using tRNAscan-SE 1.21 (11) to find tRNA, by using RNAmmer 1.2 (8) to search rRNA, and by using Tandem Repeats Finder 4.04 (1) to find tandem repeat sequence. In addition, the contigs were searched against the KEGG (5), Pfam (2), COGs (16), and NCBI NR databases to annotate the genome.

The draft genome includes 8,581,920 bases and contains 8,302 predicted coding sequences (CDSs), with a G+C content of 67.94%. There are single-copy genes predicted for 16S and 23S rRNA, duplicated genes predicted for 5S rRNA, and 50 copies predicted for tRNAs. An estimated 82.8% of nucleotides are predicted to encode proteins. We also predicted 615 genes related to regulatory function and identified 913 tandem repeat regions. The CDSs annotated by COGs can be classified into 23 COG (clusters of orthologous groups) categories and 1,845 COGs, and 2962 CDSs can be annotated in the KEGG orthology system.

The following are present in the genome sequence: genes in the spinosyn cluster (*spnA*, *spnB*, *spnC*, *spnD*, and *spnE*), the PKS-encoding genes (17); the 4′-, 2′-, and 3′-*O*-methyltransferase genes (*spnH*, *spnI*, and *spnK*, respectively [7]); *spnJ*, which has catalytic function during the synthesis of the precursor of the tricyclic nucleus of spinosyns (6); and *spnQ*, which is responsible for the biosynthesis of D-forosamine (4).

Nucleotide sequence accession number. This whole genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number AEYC00000000.

This work was supported by the National Natural Science Foundation of China (NFSC) general projects grant 30770060, and the National High Technology Research and Development Program of China (863 Program: 2007AA021301).

REFERENCES

1. **Benson, G.** 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. **27:**573–580.
2. **Finn, R. D., et al.** 2010. The Pfam protein families database. Nucleic Acids Res. **38:**D211–D222.
3. **Grouse, G. D., and T. C. Sparks.** 1998. Naturally derived materials as products and leads for insect control: the spinosyns. Pesticides and the future: minimizing chronic exposure of humans and the environment 133.
4. **Hong, L., Z. Zhao, and H. W. Liu.** 2006. Characterization of SpnQ from the

* Corresponding author. Mailing address: Microbial Genome Research Center, CAS Key Lab of Pathogenic Microbiology and Immunology, Institute of Microbiology, CAS, NO.1 West Beichen Road, Chaoyang District, Beijing 100101, P. R. China. Phone: 86-10-64807362. Fax: 86-10-64807438. E-mail: zhubaoli@im.ac.cn.

▽ Published ahead of print on 8 April 2011.

spinosyn biosynthetic pathway of Saccharopolyspora spinosa: mechanistic and evolutionary implications for C-3 deoxygenation in deoxysugar biosynthesis. J. Am. Chem. Soc. **128:**14262–14263.
5. **Kanehisa, M., S. Goto, S. Kawashima, Y. Okuno, and M. Hattori.** 2004. The KEGG resource for deciphering the genome. Nucleic Acids Res. **32:**D277–D280.
6. **Kim, H. J., R. Pongdee, Q. Wu, L. Hong, and H. W. Liu.** 2007. The biosynthesis of spinosyn in Saccharopolyspora spinosa: synthesis of the cross-bridging precursor and identification of the function of SpnJ. J. Am. Chem. Soc. **129:**14582–14584.
7. **Kim, H. J., et al.** 2010. Biosynthesis of spinosyn in Saccharopolyspora spinosa: synthesis of permethylated rhamnose and characterization of the functions of SpnH, Spnl, and SpnK. J. Am. Chem. Soc. **132:**2901–2903.
8. **Lagesen, K., et al.** 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. **35:**3100–3108.
9. **Li, R., et al.** 2010. De novo assembly of human genomes with massively parallel short read sequencing. Genome Res. **20:**265–272.
10. **Liolios, K., et al.** 2010. The Genomes On Line Database (GOLD) in 2009: status of genomic and metagenomic projects and their associated metadata. Nucleic Acids Res. **38:**D346–D354.
11. **Lowe, T. M., and S. R. Eddy.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25:**955–964.
12. **Mertz, F. P., and R. C. Yao.** 1990. Saccharopolyspora spinosa sp. nov. isolated from soil collected in a sugar mill rum still. Int. J. Syst. Bacteriol. **40:**34–39.
13. **Noguchi, H., T. Taniguchi, and T. Itoh.** 2008. MetaGeneAnnotator: detecting species-specific patterns of ribosomal binding site for precise gene prediction in anonymous prokaryotic and phage genomes. DNA Res. **15:**387–396.
14. **Sparks, T. C., et al.** 1998. Biological activity of the spinosyns, new fermentation derived insect control agents, on tobacco budworm (Lepidoptera: Noctuidae) larvae. J. Econ. Entomol. **91:**1277–1283.
15. **Sparks, T. C., et al.** 1999. Fermentation-derived insect control agents. Biopesticides Use Delivery 171–188.
16. **Tatusov, R. L., M. Y. Galperin, D. A. Natale, and E. V. Koonin.** 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. **28:**33–36.
17. **Waldron, C., et al.** 2001. Cloning and analysis of the spinosad biosynthetic gene cluster of Saccharopolyspora spinosa. Chem. Biol. **8:**487–499.